The t-test

Some problems with the z-test

Unknown SD "adds uncertainty"
(Lower draws also problem)

when to use t-test

1) Draws from box

2) $n \leq 25$ ($n :=$ #draws)

3) SD unknown

4) Data close to normal

The t-test mechanics

$$t = \frac{\text{value} - EV}{SE^*}$$

then for p-value use $p(t, df)$

$df = n-1$

$SE^* = SE$ using SD^*

$$SD^* = \sqrt{\frac{n}{n-1}} \cdot SD_{sample}$$

NOTE that in class

$$SD = \sqrt{\frac{1}{n}\sum_{x}(x-avg)^2}$$

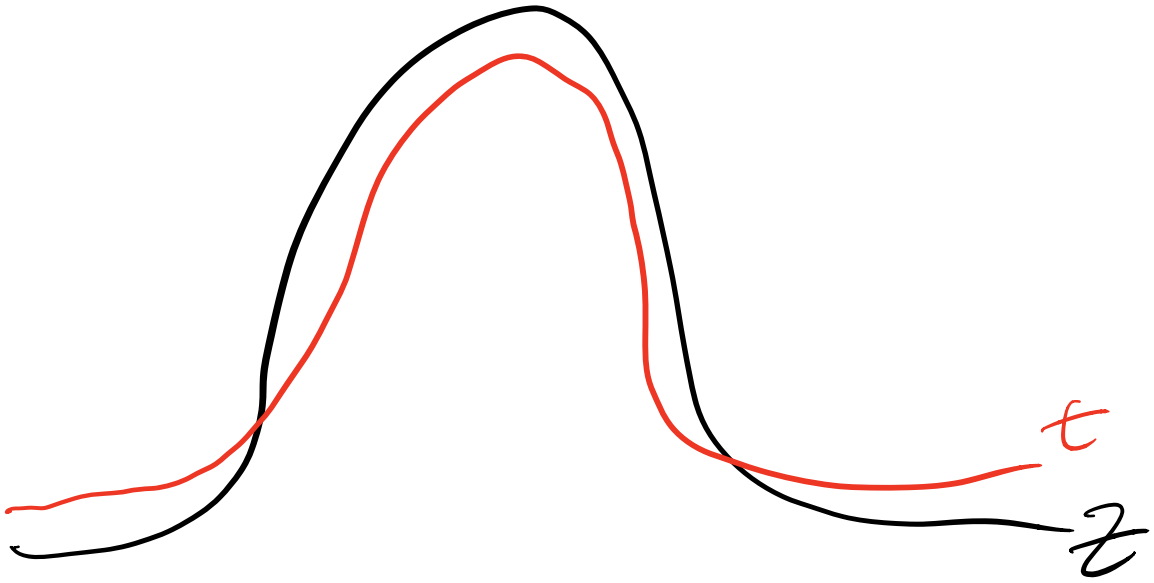
R's Sd function

$$Sd = \sqrt{\frac{1}{n-1}\sum_{x}(x-avg)^2}$$

Why do we use t?

adding uncertainty from estimating SD

"thicker" tails in t



more density in tails

$\Rightarrow$ more extreme events are more probable.

What follows is completely extra if you don't understand don't worry

Context

I said as a correction only use

$$\text{Stat} \pm t_{1-\frac{\alpha}{2}} \cdot SE$$

for $(1-\alpha) \cdot 100\%$ CI

if a lot of draws

or if SD known